CSE 203B Week 3 Discussion

Enming Guo

2025.01.23

Reminders

- Course website: https://cseweb.ucsd.edu//classes/wi25/cse203B-a/
- HW2 due on Jan 23 (Thursday) 11:50 PM
- Late policy for homework: Piazza note @19

Contents

Support Vector Machine

Input:

$$(x_i, y_i), i = 1, \dots, m;\ x_i \in \mathbb{R}^n, y_i \in \{1, -1\} \text{(labels, two classes)}$$

Hard Margin SVM Objective:

Find$(a, b), where\ a \in \mathbb{R}^n$,$b \in \mathbb{R}$, such that:

$$min\|a\|^2$$

Subject to:

$$y_i(a^T x_i - b) \geq 1, \forall 1 \leq i \leq m$$

The hyperplane is represented by $a \in \mathbb{R}^n$and $b \in \mathbb{R}$, where a is the normal vector, and is the b bias term.

Support Vector Machine

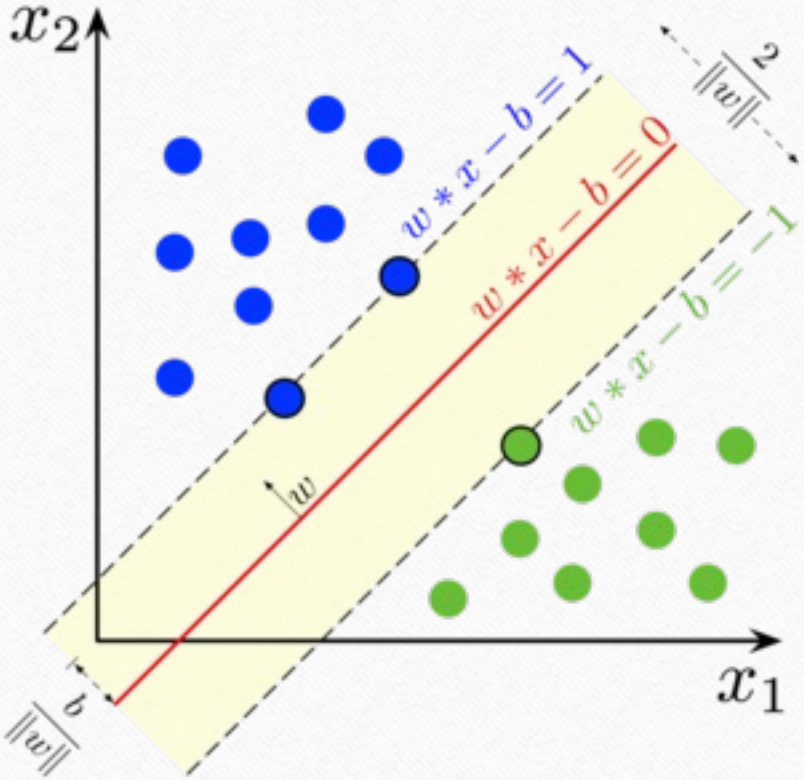
SVM seeks a hyperplane that satisfies the following two conditions:

1.Correct Classification: For all samples i, the following condition must hold:

$$y_i(a^T x_i - b) \geq 1$$

This is the classification constraint, ensuring that positive samples $(y_i = 1)$lie on one side of the hyperplane, and negative samples $(y_i = -1)$lie on the other.

2.Maximizing the Margin: The margin is defined as $\frac{2}{\|a\|}$. SVM maximizes the margin by minimizing $\|a\|^2$.



Support Vector Machine

Input:

$$(x_i, y_i), i = 1, \dots, m;\ x_i \in \mathbb{R}^n, y_i \in \{1, -1\}(\text{labels, two classes})$$

Soft Margin 1: Find$(a, b), where\ a \in \mathbb{R}^n,\ \ b \in \mathbb{R},$

$$\min_{a,b,\lambda} \lambda\|a\|^2 + \frac{1}{m}\sum Max\big(0,1 - y_i(a^T x_i - b)\big)\ \ a \in \mathbb{R}^n, \mathrm{b} \in \mathbb{R}, \lambda > 0$$

Soft Margin 2: Find$(a, b), where\ a \in \mathbb{R}^n,\ \ b \in \mathbb{R}, c_i \in R_+, i = 1, \dots, m$

$$\min_{a,b} \|a\|^2 + C\left[\frac{1}{m}\sum Max\big(0,1 - y_i(a^T x_i - b)\big)\right] a \in \mathbb{R}^n, \mathrm{b} \in \mathbb{R}, C > 0$$

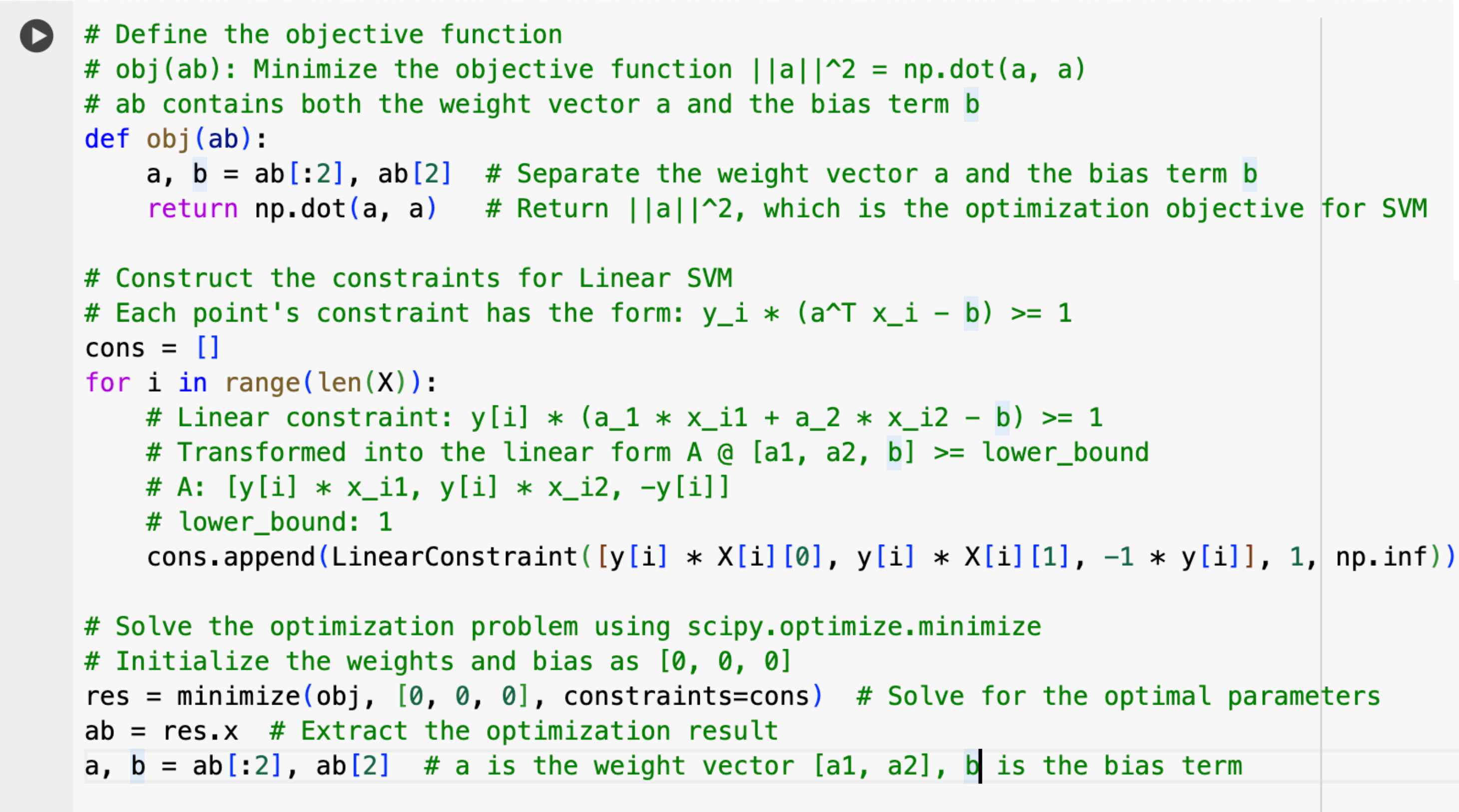
By deconstructing the hinge loss, the following formulation is obtained:

$$\min_{a,b,\zeta} \|a\|^2 + C\sum_{i=1}^{m} \xi_i\ (slack\ variables)$$

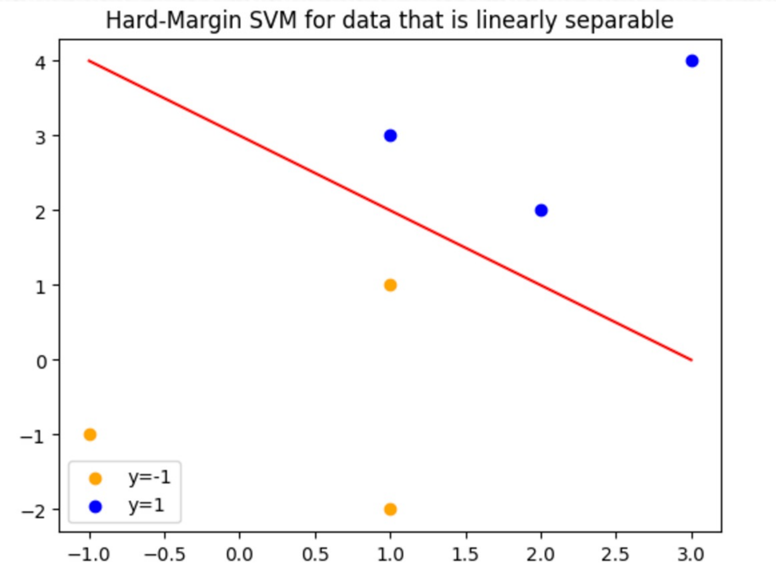
$$s.t\ y_i(a^T x - b) \geq 1 - \xi_i, \xi_i\ \geq 0$$

Support Vector Machine

Code Solution for Hard Margin SVM

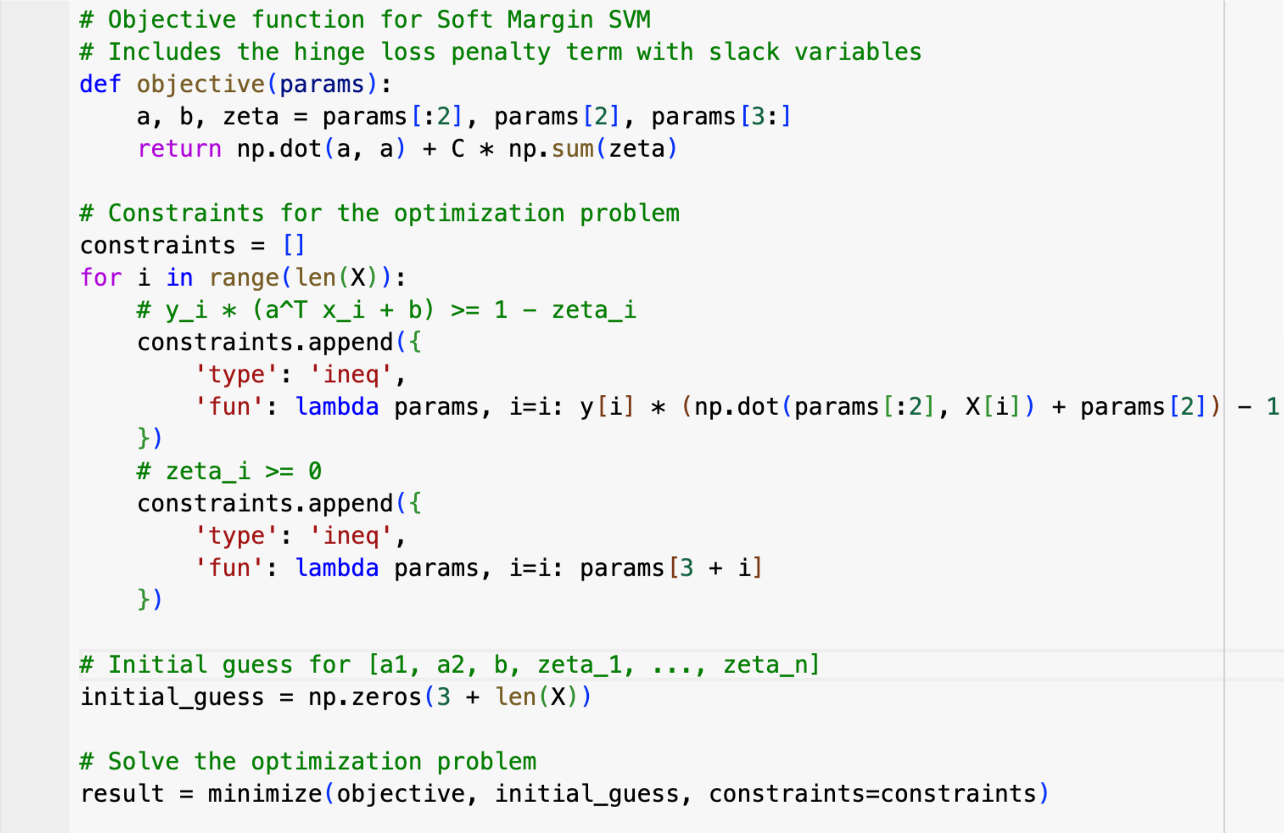


```
# Define the objective function
# obj(ab): Minimize the objective function ||a||^2 = np.dot(a, a)
# ab contains both the weight vector a and the bias term b
def obj(ab):
    a, b = ab[:2], ab[2]  # Separate the weight vector a and the bias term b
    return np.dot(a, a)   # Return ||a||^2, which is the optimization objective for SVM

# Construct the constraints for Linear SVM
# Each point's constraint has the form: y_i ∗ (a^T x_i − b) >= 1
cons = []
for i in range(len(X)):
    # Linear constraint: y[i] ∗ (a_1 ∗ x_i1 + a_2 ∗ x_i2 − b) >= 1
    # Transformed into the linear form A @ [a1, a2, b] >= lower_bound
    # A: [y[i] ∗ x_i1, y[i] ∗ x_i2, −y[i]]
    # lower_bound: 1
    cons.append(LinearConstraint([y[i] ∗ X[i][0], y[i] ∗ X[i][1], −1 ∗ y[i]], 1, np.inf))

# Solve the optimization problem using scipy.optimize.minimize
# Initialize the weights and bias as [0, 0, 0]
res = minimize(obj, [0, 0, 0], constraints=cons)  # Solve for the optimal parameters
ab = res.x  # Extract the optimization result
a, b = ab[:2], ab[2]  # a is the weight vector [a1, a2], b is the bias term
```

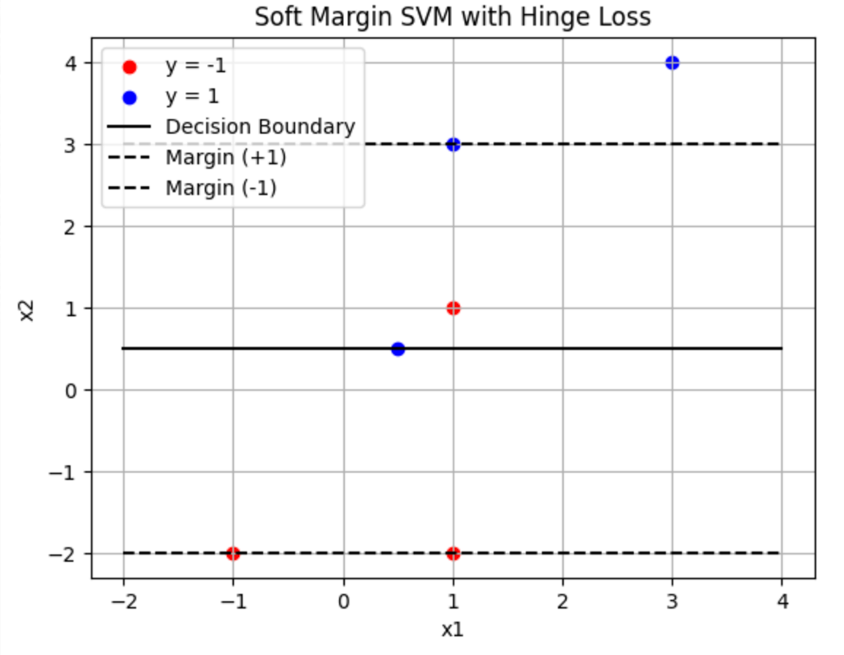


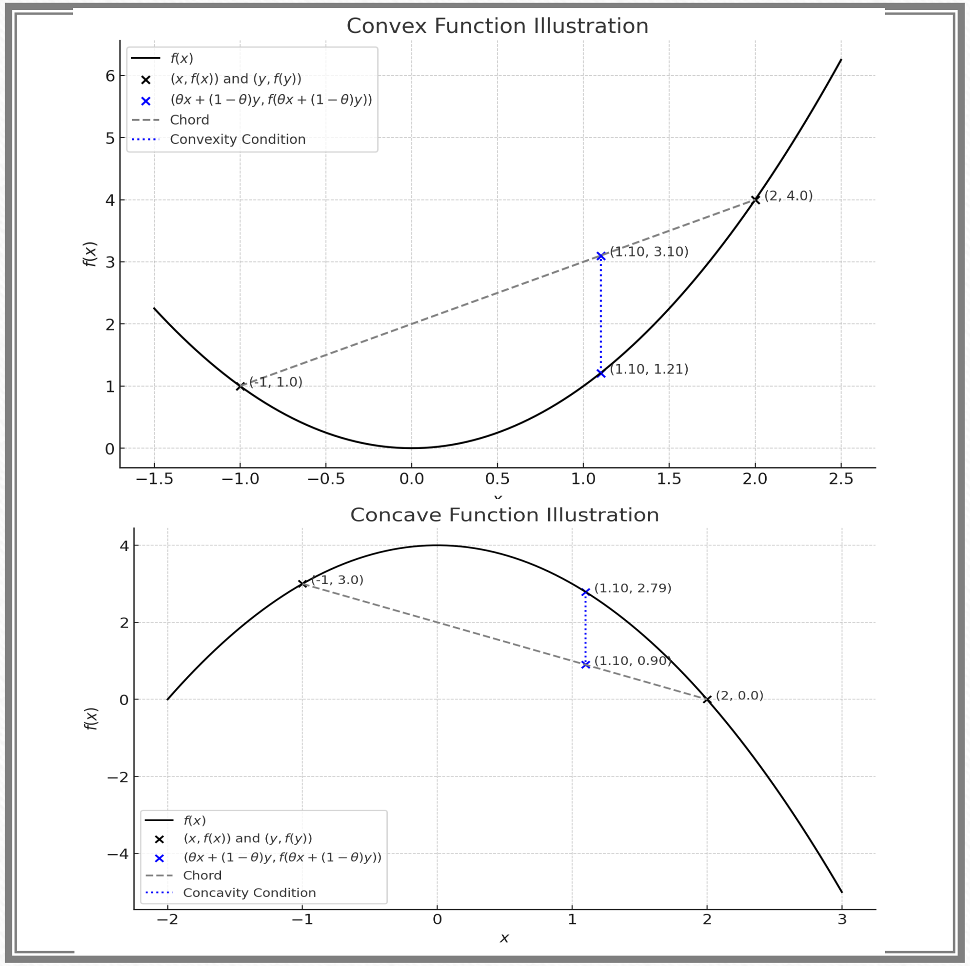
Support Vector Machine

Code Solution for Soft Margin SVM



```
# Objective function for Soft Margin SVM
# Includes the hinge loss penalty term with slack variables
def objective(params):
    a, b, zeta = params[:2], params[2], params[3:]
    return np.dot(a, a) + C * np.sum(zeta)

# Constraints for the optimization problem
constraints = []
for i in range(len(X)):
    # y_i * (a^T x_i + b) >= 1 - zeta_i
    constraints.append({
        'type': 'ineq',
        'fun': lambda params, i=i: y[i] * (np.dot(params[:2], X[i]) + params[2]) - 1 + params[3 + i]
    })
    # zeta_i >= 0
    constraints.append({
        'type': 'ineq',
        'fun': lambda params, i=i: params[3 + i]
    })

# Initial guess for [a1, a2, b, zeta_1, ..., zeta_n]
initial_guess = np.zeros(3 + len(X))

# Solve the optimization problem
result = minimize(objective, initial_guess, constraints=constraints)
```





Convex Functions

- A function $f\colon \mathbb{R}^n \to \mathbb{R}$ is **convex** if:

Its domain ($\boldsymbol{dom\ f}$) is convex.

A convex domain means that for any two points x, y in the domain, the line segment joining x and y lies entirely within the domain.

The following inequality holds for all $\boldsymbol{x, y \in dom\ f\ and\ \theta \in [0,1]}$:

$$\boldsymbol{f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta)f(y)}$$

- A function f is concave if $-f$ is convex

$$f(\theta x + (1-\theta)y) \ge \theta f(x) + (1-\theta)f(y)$$

$$-f(\theta x + (1-\theta)y) \le -\theta f(x) - (1-\theta)f(y)$$

- f is convex if and only if for all $x \in dom\ f$ and all directions v, the function $g(t) = f(x + tv)$ is convex on its domain $\{t | x + tv \in dom f\}$

Convex Functions

Prove that the following function is convex using the **Directional Convexity Definition**:

$$f(x) = \|Ax - b\|_2^2, x \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$$

Convex Functions ---- First Order Conditions

Suppose f is differentiable, i.e. $dom\ f$ is open and ∇f exists for all $x \in dom\ f$, then f is convex if and only if $dom\ f$ is convex, and ,

$$f(y) \geq f(x) + \nabla f(x)^T (y - x)\ \ \forall x, y \in dom\ f$$

$\nabla f(x)$: $The\ gradient\ of\ f\ at\ x$
$\nabla f(x)^T (y - x)$: The first order of taylor approximation of $f(y)$ around x

Convex Functions ---- First Order Conditions

Prove that the following function is convex using the **First Order Condition**:

$$f(x) = \|x\|_2^2, x \in \mathbb{R}^n$$

Convex Functions ---- Second Order Conditions

Suppose f is twice differentiable, i.e, $dom\ f$ is open and the Hessian $\nabla^2 f$ exists for all $x \in dom\ f$, then f is convex if and only if $dom\ f$ is convex, and

$$\nabla^2 f(x) \succcurlyeq 0, \forall x \in dom\ f$$

Hessian Matrix is positive semidefinite

- ○ All eigenvalues ≥ 0
- ○ $x^T \nabla^2 f(x) x \geq 0, \forall x$

Convex Functions ---- Second Order Conditions

Prove that the following function is convex using the **Second Order Condition**:

$$f(x,y) = \frac{x^2}{y} \quad dom\, f = \mathbb{R} \times \mathbb{R}_{++} = \{(x,y) \in \mathbb{R}^2 | y > 0\}$$

Convex Functions Example

Convex Functions

- Affine functions, $Ax + b$
- Exponential, $e^{ax}, x \in R, a \in R$
- Powers, x^a when $a \geq 1 \; or \; a \leq 0, x \in \mathbb{R}_{++}$
- Norms, $\|x\|_p, x \in R^n, p \geq 1$
- Max functions, $\max(x)\,, x \in R^n$
- Log-Sum-Exponential Function,

$$f(x) = \log\left(\sum_{i=1}^{n} e^{x_i}\right), x \in \mathbb{R}^n$$

Concave Functions

- Affine functions, $Ax + b$
- Logarithm, $logx$
- Powers, x^a when $0 \leq a \leq 1\,, x \in \mathrm{R}_{++}$
- Logarithm of Determinant
  $f(X) = logdet(X), X \succ 0$ (Positive Definite Matrix)

Convex Functions ---- Operations that preserve convexity

- **Non-negative Weighted Sums**

  if $f_1(x)$ and $f_2(x)$ are convex functions, their weighted sum is also convex if the weights are non-negative

  $$f(x) = w_1 f_1(x) + w_2 f_2(x), w_1, w_2 \geq 0$$

- **Composition of Functions**

  $f(x) = h(g(x))$, where $g(x)$ is the inner function and $h(y)$ is the outer function.

  Conditions for Convexity:

  1.If $h(y)$ is convex and non-decreasing, and $g(x)$ is convex, then $f(x)$ is convex.

  2.If $h(y)$ is convex and non-decreasing, and $g(x)$ is concave, then $f(x)$ is concave.

  Conditions for Concavity:

  1.If $h(y)$ is concave and non-decreasing, and $g(x)$ is concave, then $f(x)$ is concave.

  2.If $h(y)$ is concave and non-decreasing, and $g(x)$ is convex, then $f(x)$ is convex.

- **Pointwise Maximum and Supremum**

  $$f(x) = max\{f_1(x), f_2(x)\}$$

  $$f(x) = sup_y g(x, y)\ [g(x, y)\ is\ convex\ in\ x\ for\ each\ fixed\ y]$$

- **Partial Minimization**

  $$f(x) = inf_y g(x, y)$$

Convex Functions ---- Proof of the convexity or concavity

- By definition
- Reduce multivariate function to 1-dimensional $g(t)$
- First order conditions
- Second order conditions
- Properties

…